the cells expressing surface antigen. Rosettes can be formed between antigen expressing cells and erythrocytes bearing covalently coupled antibody to the relevant antigen. These are readily purified by unit gravity sedimentation. Panning of the cell population over petri dishes bearing immobilized monoclonal antibody specific for the relevant antigen can also be used to remove unwanted cells.

In the high throughput assays of the invention, it is possible to screen up to several thousand different experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) variants in a single day. For example, each well of a microtiter plate can be used to run a separate assay, or, if concentration or incubation time effects are to be observed, every 5 -10 wells can test a single variant. Thus, a single standard microtiter plate can assay about 100 (e.g., 96) reactions. If 1536 well plates are used, then a single plate can easily assay from about 100 to about 1500 different reactions. It is possible to assay several different plates per day; assay screens for up to about 6,000-20,000 different assays (i.e., involving different nucleic acids, encoded proteins, concentrations, etc.) is possible using the integrated systems of the invention. More recently, microfluidic approaches to reagent manipulation have been developed, e.g., by Caliper Technologies (Palo Alto, CA).

In one aspect, library members, e.g., cells, viral plaques, or the like, are separated on solid media to produce individual colonies (or plaques). Using an automated colony picker (e.g., the Q-bot, Genetix, U.K.), colonies or plaques are identified, picked, and up to 10,000 different mutants inoculated into 96 well microtiter dishes, optionally containing glass balls in the wells to prevent aggregation. The Q-bot does not pick an entire colony but rather inserts a pin through the center of the colony and exits with a small sampling of cells (or viruses in plaque applications). The time the pin is in the colony, the number of dips to inoculate the culture medium, and the time the pin is in that medium each effect inoculum size, and each can be controlled and optimized. The uniform process of the Q-bot decreases human handling error and increases the rate of establishing cultures (roughly 10,000/4 hours). These cultures are then shaken in a temperature and humidity controlled incubator. The glass balls in the microtiter plates act to promote uniform aeration of cells dispersal of cells, or the like, similar to the blades of a fermentor. Clones from cultures of interest can be cloned by limiting dilution. Plaques or cells constituting libraries can also be screened

directly for production of proteins, either by detecting hybridization, protein activity, protein binding to antibodies, or the like.

The ability to detect a subtle increase in the performance of a experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) library member over that of a parent strain relies on the sensitivity of the assay. The chance of finding the organisms having an improvement in ability to induce an immune response is increased by the number of individual mutants that can be screened by the assay. To increase the chances of identifying a pool of sufficient size, a prescreen that increases the number of mutants processed by I 0-fold can be used. The goal of the prescreen will be to quickly identify mutants having equal or better product titers than the parent strain(s) and to move only these mutants forward to liquid cell culture for subsequent analysis.

A number of well known robotic systems have also been developed for solution phase chemistries useful in assay systems. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.) which mimic the manual synthetic operations performed by a scientist. Any of the above devices are suitable for use with the present invention, e.g., for high- throughput screening of molecules encoded by codon-altered nucleic acids. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein with reference to the integrated system will be apparent to persons skilled in the relevant art.

High throughput screening systems are commercially available (see, e.g., Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate entire procedures including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization.

The manufacturers of such systems provide detailed protocols the various high throughput. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the

like. Microfluidic approaches to reagent manipulation have also been developed, e.g., by Caliper Technologies (Palo Alto, CA).

Optical images viewed (and, optionally, recorded) by a camera or other recording device (e.g., a photodiode and data storage device) are optionally further processed in any of the embodiments herein, e.g., by digitizing the image and/or storing and analyzing the image on a computer. As noted above, in some applications, the signals resulting from assays are florescent, making optical detection approaches appropriate in these instances. A variety of commercially available peripheral equipment and software is available for digitizing, storing and analyzing a digitized video or digitized optical image, e.g., using PC (Intel x86 or Pentium chip- compatible DOS, OS2 WINDOWS, WINDOWS NT or VIMOWS95 based machines), MACINTOSH, or LTNIX based (e.g., SLJN work station) computers.

One conventional system carries light from the assay device to a cooled charge-coupled device (CCD) camera, in common use in the art. A CCD camera includes an array of picture elements (pixels). The light from the specimen is imaged on the CCD. Particular pixels corresponding to regions of the specimen (e.g., individual hybridization sites on an array of biological polymers) are sampled to obtain light intensity readings for each position. Multiple pixels are processed in parallel to increase speed. The apparatus and methods of the invention are easily used for viewing any sample, e.g., by fluorescent or dark field microscopic techniques.

Integrated systems for analysis in the present invention typically include a digital computer with high-throughput liquid control software, image analysis software, data interpretation software, a robotic liquid control armature for transferring solutions from a source to a destination operably linked to the digital computer, an input device (e.g., a computer keyboard) for entering data to the digital computer to control high throughput liquid transfer by the robotic liquid control armature and, optionally, an image scanner for digitizing label signals from labeled assay component. The image scanner interfaces with the image analysis software to provide a measurement of optical intensity. Typically, the intensity measurement is interpreted by the data interpretation software to show whether the optimized recombinant antigenic polypeptide products are produced.

ANTIGEN LIBRARY IMMUNIZATION

In one embodiment, antigen library immunization (ALI) is used to identify optimized recombinant antigens that have improved immunogenicity. ALI involves introduction of the

library of recombinant antigen-encoding nucleic acids, or the recombinant antigens encoded by the experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) nucleic acids, into a test animal. The animals are then subjected to *in vivo* challenge using live pathogens. Neutralizing antibodies and cross-protective immune responses are studied after immunization with the entire libraries, pools and/or individual antigen variants.

Methods of immunizing test animals are well known to those of skill in the art. In one embodiments, test animals are immunized twice or three times at two week intervals. One week after the last immunization, the animals are challenged with live pathogens (or mixtures of pathogens), and the survival and symptoms of the animals is followed. Immunizations using test animal challenge are described in, for example, Roggenkamp et al. (1997) Infect. Immun. 65: 446; Woody et al. (1997) Vaccine 2: 133; Agren et al. (1997) J Immunol. 158: 3936; Konishi et al. (1992) Virology 190: 454; Kinney et al. (1988) J Virol. 62: 4697; Iacono-Connors et al. (1996) Virus Res. 43: 125; Kochel et al. (1997) Vaccine 15: 547; and Chu et al. (1995) J Virol. 69: 6417.

The immunizations can be performed by injecting either the experimentally generated polynucleotides themselves, i.e., as a genetic vaccine, or by immunizing the animals with polypeptides encoded by the experimentally generated polynucleotides. Bacterial antigens are typically screened primarily as recombinant proteins, whereas viral antigens can be analyzed using genetic vaccinations.

To dramatically reduce the number of experiments required to identify individual antigens having improved immunogenic properties, one can use pooling and deconvolution, as diagrammed herein. Pools of recombinant nucleic acids, or polypeptides encoded by the recombinant nucleic acids, are used to immunize test animals. Those pools that result in protection against pathogen challenge are then subdivided and subjected to additional analysis. The high throughput *in vitro* approaches described above can be used to identify the best candidate sequences for the *in vivo* studies.

The challenge models that can be used to screen for protective antigens include pathogen and toxin models, such as Yersinia bacteria, bacterial toxins (such as Staphylococcal and Streptococcal enterotoxins, E. coli/V. cholerae enterotoxins), Venezuelan equine encephalitis virus (VEE), Flaviviruses (Japanese encephalitis virus, Tick-borne encephalitis virus, Dengue virus), Hantaan virus, Herpes simplex, influenza virus (e.g.,

Influenza A virus), Vesicular Steatites Virus, Pseudomonas aeruginosa, Salmonella typhimurium, Escherichia coli, Klebsiella pneumoniae, Toxoplasma gondii, Plasmodium yoeliii, Herpes simplex, influenza virus (e.g., Influenza A virus), and Vesicular Steatites Virus. However, the test animals can also be challenged with tumor cells to enable screening of antigens that efficiently protect against malignancies. Individual experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) antigens or pools of antigens are introduced into the animals intradermally, intramuscularly, intravenously, intratracheally, anally, vaginally, orally, or intraperitoneally and antigens that can prevent the disease are chosen, when desired, for further rounds of reassembly (optionally in combination with other directed evolution methods described herein) and selection. Eventually, the most potent antigens, based on *in vivo* data in test animals and comparative *in vitro* studies in animals and man, are chosen for human trials, and their capacity to prevent and treat human diseases is investigated.

In some embodiments, antigen library immunization and pooling of individual clones is used to immunize against a pathogen strain that was not included in the sequences that were used to generate the library. The level of crossprotection provided by different strains of a given pathogen can significantly. However, homologous titer is always higher than heterologous titer. Pooling and deconvolution is especially efficient in models where minimal protection is provided by the wild-type antigens used as starting material for reassembly (optionally in combination with other directed evolution methods described herein). This approach can be taken, for example, when evolving the V-antigen of Yersinae or Hantaan virus glycoproteins.

In some embodiments, the desired screening involves analysis of the immune response based on immunological assays known to those skilled in the art. Typically, the test animals are first immunized and blood or tissue samples are collected for example one to two weeks after the last immunization. These studies enable one to one can measure immune parameters that correlate to protective immunity, such as induction of specific antibodies (particularly IgG) and induction of specific T lymphocyte responses, in addition to determining whether an antigen or pools of antigens provides protective immunity.

Spleen cells or peripheral blood mononuclear cells can be isolated from immunized test animals and measured for the presence of antigen-specific T cells and induction of

cytokine synthesis. ELISA, ELISPOT and cytoplasmic cytokine staining, combined with flow cytometry, can provide such information on a single-cell level.

Common immunological tests that can be used to identify the efficacy of immunization include antibody measurements, neutralization assays and analysis of activation levels or frequencies of antigen presenting cells or lymphocytes that are specific for the antigen or pathogen. The test animals that can be used in such studies include, but are not limited to, mice, rats, guinea pigs, hamsters, rabbits, cats, dogs, pigs and monkeys.

Monkey is a particularly useful test animal because the MHC molecules of monkeys and humans are very similar. Virus neutralization assays are useful for detection of antibodies that not only specifically bind to the pathogen, but also neutralize the function of the virus. These assays are typically based on detection of antibodies in the sera of immunized animal and analysis of these antibodies for their capacity to inhibit viral growth in tissue culture cells. Such assays are known to those skilled in the art. One example of a virus neutralization assay is described by Dolin R (J. Infect. Dis. 1995, 172:1175-83). Virus neutralization assays provide means to screen for antigens that also provide protective immunity.

In some embodiments, experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) antigens are screened for their capacity to induce T cell activation *in vivo*. More specifically, peripheral blood mononuclear cells or spleen cells from injected mice can be isolated and the capacity of cytotoxic T lymphocytes to lyse infected, autologous target cells is studied. The spleen cells can be reactivated with the specific antigen *in vitro*. In addition, T helper cell activation and differentiation is analyzed by measuring cell proliferation or production of T_H (IL-2 and IFN-γ) and T_H2 (IL- 4 and IL-5) cytokines by ELISA and directly in CD4+ T cells by cytoplasmic cytokine staining and flow cytometry. Based on the cytokine production profile, one can also screen for alterations in the capacity of the antigens to direct T_H1/ T_H2 differentiation (as evidenced, for example, by changes in ratios of IL-4/ IFN-γ, IL-4/IL-2, IL-5/ IFN-γ, IL-5/IL-2, IL- 13/ IFN-γ, IL- I 3/IL-2). The analysis of the T cell activation induced by the antigen variants is a very useful screening method, because potent activation of specific T cells *in vivo* correlates to induction of protective immunity.

The frequency of antigen-specific CD8+ T cells *in vivo* can also be directly analyzed using tetramers of MHC class I molecules expressing specific peptides derived from the

corresponding pathogen antigens (Ogg and McMichael, Curr. Opin. Immunol. 1998, 10:393-6; Altman et al., Science 1996, 274:94-6). The binding of the tetramers can be detected using flow cytometry, and will provide information about the efficacy of the experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) antigens to induce activation of specific T cells. For example, flow cytometry and tetramer stainings provide an efficient method of identifying T cells that are specific to a given antigen or peptide. Another method involves panning using plates coated with tetramers with the specific peptides. This method allows large numbers of cells to be handled in a short time, but the method only selects for highest expression levels. The higher the frequency of antigen-specific T cells *in vivo* is, the more efficient the immunization has been, enabling identification of the antigen variants that have the most potent capacity to induce protective immune responses. These studies are particularly useful when conducted in monkeys, or other primates, because the MHC class I molecules of humans mimic those of other primates more closely than those of mice.

Measurement of the activation of antigen presenting cells (APC) in response to immunization by antigen variants is another useful screening method. Induction of APC activation can be detected based on changes in surface expression levels of activation antigens, such as 137-1 (CD80). 137-2 (CD86), MHC class I and 11, CD14, CD23, and Fc receptors, and the like.

Experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) cancer antigens that induce cytotoxic T cells that have the capacity to kill cancer cells can be identified by measuring the capacity of T cells derived from immunized animals to kill cancer cells *in vitro*. Typically the cancer cells are first labeled with radioactive isotopes and the release of radioactivity is an indication of tumor cell killing after incubation in the presence of T cells from immunized animals. Such cytotoxicity assays are known in the art.

An indication of the efficacy of an antigen to activate T cells specific for, for example, cancer antigens, allergens or autoantigens, is also the degree of skin inflammation when the antigen is injected into the skin of a patient or test animal. Strong inflammation is correlated with strong activation of antigen-specific T cells. Improved activation of tumor-specific T cells may lead to enhanced killing of the tumors. In case of autoantigens, one can add immunomodulators that skew the responses towards T_H2, whereas in the case of

allergens a T_H1 response is desired. Skin biopsies can be taken, enabling detailed studies of the type of immune response that occurs at the sites of each injection (in mice and monkeys large numbers of injections/antigens can be analyzed). Such studies include detection of changes in expression of cytokines, chemokines, accessory molecules, and the like, by cells upon injection of the antigen into the skin.

To screen for antigens that have optimal capacity to activate antigen-specific T cells, peripheral blood mononuclear cells from previously infected or immunized humans individuals can be used. This is a particularly useful method, because the MHC molecules that will present the antigenic peptides are human MHC molecules. Peripheral blood mononuclear cells or purified professional antigen-presenting cells (APCs) can be isolated from previously vaccinated or infected individuals or from patients with acute infection with the pathogen of interest. Because these individuals have increased frequencies of pathogen-specific T cells in circulation, antigens expressed in PBMCs or purified APCs of these individuals will induce proliferation and cytokine production by antigen- specific $CD4^+$ and $CD8^+$ T cells. Thus, antigens that simultaneously harbor epitopes from several antigens can be recognized by their capacity to stimulate T cells from various patients infected or immunized with different pathogen antigens, cancer antigens, autoantigens or allergens. One buffy coat derived from a blood donor contains lymphocytes from 0.5 liters of blood, and up to 10^4 PBMC can be obtained, enabling very large screening experiments using T cells from one donor.

When healthy vaccinated individuals (lab volunteers) are studied, one can make EBV-transformed B cell lines from these individuals. These cell lines can be used as antigen presenting cells in subsequent experiments using blood from the same donor; this reduces interassay and donor-to-donor variation. In addition, one can make antigen-specific T cell clones, after which antigen variants are introduced to EBV transformed B cells. The efficiency with which the transformed B cells induce proliferation of the specific T cell clones is then studied. When working with specific T cell clones, the proliferation and cytokine synthesis responses are significantly higher than when using total PBMCs, because the frequency of antigen-specific T cells among PBMC is very low.

CTL epitopes can be presented by most cells types since the class I major histocompatibility complex (MHC) surface glycoproteins are widely expressed. Therefore, transfection of cells in culture by libraries of experimentally evolved (e.g. by polynucleotide

reassembly &/or polynucleotide site-saturation mutagenesis) antigen sequences in appropriate expression vectors can lead to class I epitope presentation. If specific CTLs directed to a given epitope have been isolated from an individual, then the co-culture of the transfected presenting cells and the CTLs can lead to release by the CTLs of cytokines, such as IL-2, IFN-γ, or TNF, if the epitope is presented. Higher amounts of released TNF will correspond to more efficient processing and presentation of the class I epitope from the experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis), evolved sequence. Experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) antigens that induce cytotoxic T cells that have the capacity to kill infected cells can also be identified by measuring the capacity of T cells derived from immunized animals to kill infected cells *in vitro*. Typically the target cells are first labeled with radioactive isotopes and the release of radioactivity is an indication of target cell killing after incubation in the presence of T cells from immunized animals. Such cytotoxicity assays are known in the art.

A second method for identifying optimized CTL epitopes does not require the isolation of CTLs reacting with the epitope. In this approach, cells expressing class I MHC surface glycoproteins are transfected with the library of evolved sequences as above. After suitable incubation to allow for processing and presentation, a detergent soluble extract is prepared from each cell culture and after a partial purification of the MHC-epitope complex (perhaps optional) the products are submitted to mass spectrometry (Henderson et al. (1993) Proc. Nat'l. Acad. Sci. USA 90: 10275-10279). Since the sequence is known of the epitope whose presentation to be increased, one can calibrate the mass spectrogram to identify this peptide. In addition, a cellular protein can be used for internal calibration to obtain a quantitative result; the cellular protein used for internal calibration could be the MHC molecule itself. Thus one can measure the amount of peptide epitope bound as a proportion of the N4HC molecules.

Screeningfor Optimal Induction Of Protective Immunity

Vectors that can provide efficient, protective immunity are selected using lethal infection models to choose vectors that can prevent the disease for further rounds of reassembly (optionally in combination with other directed evolution methods described herein) and selection

To select genetic vaccine vectors that provide efficient protective immunity, one can screen the vector libraries in a test mammal using lethal infection models, such as Pseudomonas aeruginosa, Salmonella typhimurium, Escherichia coli, Klebsiella pneumoniae, Toxoplasma gondii, Plasmodium yoeliii, Herpes simplex, influenza virus (e.g., Influenza A virus), and Vesicular Steatites Virus. Pools of genetic vaccine vectors or individual vectors are introduced into the animals intradermally, intramuscularly, intravenously, intratracheally, anally, vaginally, orally, or intraperitoneally and vectors that can prevent the disease are chosen for further rounds of reassembly (optionally in combination with other directed evolution methods described herein) and selection.

Examples: anti-IL-4 mAbs or recombinant IL-12; recombinant IL-12 (advantage of latter model is that infection occurs through lung, common route of human pathogen invasion)

As an example, optimal vectors can be screened in mice infected with Leishmania major parasites. When injected into footpads of BALB/c mice, these parasites cause a progressive infection later resulting in a disseminated disease with fatal outcome, which can be prevented by anti-IL-4 mAbs or recombinant IL-12 (Chatelain et al. (1992) J. Immunol. 148: 1182-1187). Pools of plasmids can be injected intravenously, intraperitoneally or into footpads of these mice, and pools that can prevent the disease are chosen for further analysis and screened for vectors that can cure existing infections. The size of the footpad swelling can be followed visually providing simple yet precise monitoring of the disease progression. Mice can be infected intratracheally with Klebsiella pneumoniae resulting in lethal pneumonia, which can be prevented by recombinant IL-12 (Greenberger et al. (1996) J Immunol. 157: 3006-3012). The advantage of this model is that the infection occurs through the lung, which is a common route of human pathogen invasion. The vectors can be given to the lung together with the pathogen or they can be administered after symptoms are evident in order to screen for vectors that can cure established infections.

Example: Influenza- provides a way to screen for vectors that provide protection at very low quantities of DNA and/or high virus concentrations, and it also allows one to analyze the levels of antigen specific Abs and CTLs induced *in vivo*

In another example, the genetic vaccines are a mouse vaccination model for Influenza A virus. Influenza was one of the first models in which the efficacy of genetic vaccines was demonstrated (Ulmer et al. (1993) Science 259: 1745-1749). Several Influenza strains are lethal in mice providing an easy means to screen for efficacy of genetic vaccines.

For example, Influenza virus strain A/PR/8/34, which is available through the American Type Culture Collection (ATCC VR-95), causes lethal infection, but 100% survival can be obtained when the mice are immunized with and influenza hemagglutinin (HA) genetic vaccine (Deck et al. (1997) Vaccine 15: 71-78). This model provides a way to screen for vectors that provide protection at very low quantities of DNA and/or high virus concentrations, and it also allows one to analyze the levels of antigen specific Abs and CTLs induced *in vivo*.

Example: Mycobacterium tuberculosis (partial protection, requires major improvements)

The genetic vaccine vectors can also be analyzed for their capacity to provide protection against infections by Mycobacterium tuberculosis. This is an example of a situation where genetic vaccines have provided partial protection, and where major improvements are required.

Identification of candidate vectors followed by more testing

Once a number of candidate vectors has been identified, these vectors can be subjected to more detailed analysis in additional models. Testing in other infectious disease models (such as HSV, Mycoplasma pulmonis, RSV and/or rotavirus) will allow identification of vectors that are optimal in each infectious disease.

Optimal plasmids from the first round of screening are used as the starting material for the next round, the successful vectors are sequenced and the corresponding human genes are cloned into genetic vaccine vectors which are characterized *in vitro* for their capacity to induce differentiation of a desired trait.

In each case, the optimal plasmids from the first round of screening can be used as the starting material for the next round of reassembly (optionally in combination with other directed evolution methods described herein), assembly and selection. Vectors that are successful in animal models are sequenced and the corresponding human genes are cloned into genetic vaccine vectors. These vectors are then characterized *in vitro* for their capacity to induce differentiation of T_H1/ T_H2 cells, activation of T_H cells, cytotoxic T lymphocytes and monocytes/macrophages, or other desired trait. Eventually, the most potent vectors, based on *in vivo* data in mice and comparative *in vitro* studies in mice and man, are chosen for human trials, and their capacity to counteract various human infectious diseases is investigated.

Methods for measuring immune parameters that correlate to protective immunity

In addition to determining whether a vector pool provides protective immunity, one can measure immune parameters that correlate to protective immunity, such as induction of specific antibodies (particularly IgG) and induction of specific CTL responses. Spleen cells can be isolated from vaccinated mice and measured for the presence of antigen- specific T cells and induction of T_H1 cytokine synthesis profiles. ELISA and cytoplasmic cytokine staining, combined with flow cytometry, can provide such information on a single-cell level.

Screening Of Genetic Vaccine Vectors That Activate Human Antigen-Specific Lymphocyte Responses

Isolation of PBMCs or APCs to screen for vectors with optimal immunostimulatory properties for the human immune system

To screen for vectors with optimal immunostimulatory properties for the human immune system, peripheral blood mononuclear cells (PBMCs) or purified professional antigen-presenting cells (APCs) can be isolated from previously vaccinated or infected individuals or from patients with acute infection with the pathogen of interest.

Genetic vaccine vectors encoding the antigen for which the individuals have specific T cells can be transfected into PBMC and induction of T cell proliferation and cytokine synthesis can be measured; also possible to screen for spontaneous entry of genetic vaccine vector into APCs

Because these individuals have increased frequencies of pathogen- specific T cells in circulation, antigens expressed in PBMCs or purified APCs of these individuals will induce proliferation and cytokine production by antigen-specific CD4+ and CD8+ T cells. Thus, genetic vaccine vectors encoding the antigen for which the individuals have specific T cells can be transfected into PBMC of the individuals, after which induction of T cell proliferation and cytokine synthesis can be measured. Alternatively, one can screen for spontaneous entry of the genetic vaccine vector into A-PCs, thus providing a means by which to screen simultaneously for improved transfection efficiency, improved expression of antigen and improved induction of activation of specific T cells. Vectors with the most potent immunostimulatory properties can be screened based on their capacity to induce B cell proliferation and immunoglobulin synthesis. One buffy coat derived from a blood donor contains PBMC lymphocytes from 0.5 liters of blood, and up to 10^4 PBMC can be obtained, enabling very large screening experiments using T cells from one donor.

Making EBV-transformed B cell lines from healthy vaccinated individuals for subsequent experiments

When healthy vaccinated individuals (lab volunteers) are studied, one can make EBV-transformed B cell lines from these individuals. These cell lines can be used as antigen presenting cells in subsequent experiments using blood from the same donor; this reduces interassay and donor-to-donor variation). In addition, one can make antigen-specific T cell clones, after which genetic vaccines are transfected into EBV transformed B cells.

Efficiency with which the transformed B cells induce proliferation of the specific T cell clones

The efficiency with which the transformed B cells induce proliferation of the specific T cell clones is then studied. When working with specific T cell clones, the proliferation and cytokine synthesis responses are significantly higher than when using total PBMCs, because the frequency of antigen-specific T cells among PBMC is very low.

Transfection of cells in culture by libraries of experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) DNA sequences in appropriate expression vectors can lead to class I epitope presentation

CTL epitopes can be presented by most cells types since the class I major histocompatibility complex (MHC) surface glycoproteins are widely expressed. Therefore, transfection of cells in culture by libraries of experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) DNA sequences in appropriate expression vectors can lead to class I epitope presentation. If specific CTLs directed to a given epitope have been isolated from an individual, then the co-culture of the transfected presenting cells and the CTLs can lead to release by the CTLs of cytokines, such as IL-2, IFN-γ, or TNFα, if the epitope is presented. Higher amounts of released TNFα. will correspond to more efficient processing and presentation of the class I epitope from the experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis), evolved sequence.

Transfecting cells expressing class I MHC surface glycoproteins with library of evolved sequences, preparing a detergent soluble extract, performing a partial purification of the MHC-epitope complex, and then submitting the products to mass spectrometry

A second method for identifying optimized CTL epitopes does not require the isolation of CTLs reacting with the epitope. In this approach, cells expressing class I MHC

surface glycoproteins are transfected with the library of evolved sequences as above. After suitable incubation to allow for processing and presentation,. a detergent soluble extract is prepared from each cell culture and after a partial purification of the MHC-epitope complex (perhaps optional) the products are submitted to mass spectrometry (Henderson et al. (1993) Proc. Nat'l. Acad. Sci. USA 90: 10275-10279). Since the sequence is known of the epitope whose presentation to be increased, one can calibrate the mass spectrogram to identify this peptide. In addition, a cellular protein can be used for internal calibration to obtain a quantitative result; the cellular protein used for internal calibration could be the MHC molecule itself. Thus one can measure the amount of peptide epitope bound as a proportion of the MHC molecules.

SCID-Human Skin Model For Vaccination Studies

Use of mouse models in vaccine studies limited in that the MHC molecules in mice and man are substantially different, meaning that proteins and peptides that efficiently induce protective immune responses in mice do not necessarily function in humans

Successful genetic vaccinations require transfection of the target cells after injection of the vector, expression of the desired antigen, processing the antigen in antigen presenting cells, presentation of the antigenic peptides in the context of MHC molecules, recognition of the peptide/MHC complex by T cell receptors, interactions of T cells with B cells and professional APCs and induction of specific T cell and B cell responses. All these events could be differentially regulated in mouse and man. A limitation of mouse models in vaccine studies is the fact that the MHC molecules of mice and man are substantially different. Therefore, proteins and peptides that effectively induce protective immune responses in mice do not necessarily function in humans.

Mouse models can be used to study human tissues in mice *in vivo* for studies of transfection efficiency, transfer sequences, and gene expression levels

To overcome these limitations mouse models can be used to study human tissues in mice *in vivo*. Live pieces of human skin are xenotransplant onto the back of immunodeficient mice, such as SCID mice, allowing screening of the vector libraries for optimal properties in human cells *in vivo*. Recursive selection of episomal vectors provides strong selection pressure for vectors that remain episomal, yet provide high level of gene expression. These mice provide an excellent model for studies on transfection efficiency, transfer sequences and gene expression levels. In addition, antigen presenting cells (APCs) derived from these mice

can also be used to assess the level of antigens delivered to professional APCs, and to study the capacity of these cells to present antigens and induce activation of antigen-specific CD4+ and CD8+ T cells *in vitro*. Significantly, although SCID mice have severely deficient T and B cell components, antigen presenting cells (dendritic cells and monocytes) are relatively normal in these mice.

Rendering immunocompetent mice immunodeficient in order to aid transplantation of human tissue, enabling vaccine studies in human skin xenotransplanted into mice with genetically normal immune systems as well, due to the transient nature of the *in vivo* immunosuppression

In one embodiment of this model system, immunocompetent mice are rendered immunodeficient in order to enable transplantation of human tissue. For example, blocking of CD28 and CD40 pathways promotes long-term survival of allogeneic skin grafts in mice (Larsen et al. (1996) Nature 381: 434). Because the *in vivo* immunosuppression is transient, this model also enables vaccine studies in human skin xenotransplanted into mice with genetically normal immune systems. Several methods of blocking CD28- 137 interactions and CD40-CD40 ligand interactions are known to those of skill in the art, including, for example, administration of neutralizing anti-B7-1 and 137- 2 antibodies, soluble CTLA-4, a soluble form of the extracellular portion of CTLA-4, a fusion protein that includes CTLA-4 and an Fc portion of an IgG molecule, and neutralizing anti-CD40 or anti-CD40 ligand antibodies. Additional methods by which one can improve transient immunosuppression include administration of one or more of the following reagents: cyclosporin A, anti-IL-2 receptor α-chain Ab, soluble IL-2 receptor, IL- 10, and combinations thereof.

A model in which SCID-mice transplanted with human skin are injected with HLA-matched PBMC can be used to analyze vectors that provide long lasting expression *in vivo*. In this model, the vectors are injected, or topically applied, into the human skin.

If the HLA-matched PBMC injected into mice contains lymphocytes specific for the vector the transfected cells will be recognized, and eventually destroyed, by these vector-specific lymphocytes, providing the possibility to screen for vectors that efficiently escape destruction

Thereafter, HLA-matched PBMC are injected into these mice. If the PBMC contains lymphocytes specific for the vector, the transfected cells will be recognized, and eventually destroyed, by these vector-specific lymphocytes. Therefore, this model provides possibilities

to screen for vectors that efficiently escape destruction by the immune cells. It has been shown that human PBLs injected into mice with human skin transplants reject the organ, indicating that the CTLs reach the skin in this model. Obtaining HLA-matching skin and blood is possible (e.g. blood sample and skin graft from a patient undergoing skin removal due to malignancy, or blood and foreskin from the same infant).

SCIDhu mouse model: additionally, transplanting human skin allows studies on the efficacy of genetic vaccine vectors following injection to the skin

An additional model that is suitable for screening as described herein is the modified SCIDhu mouse model, in which pieces of human fetal thymus, liver and bone marrow are transplanted into SCID mice providing functional human immune system in mice (Roncarolo et al. (1996) Semin. Immunol. 8: 207). Functional human B and T cells, and APCs can be observed in these mice. When additionally human skin is transplanted, it is likely to allow studies on the efficacy of genetic vaccine vectors following injection into the skin. Cotransplantation of skin is likely to improve the model because it will provide an additional source of professional APCs.

Mouse Model For Studying The Efficiency Ofgenetic Vaccines In Transfecting Human Muscle Cells And Inducing Human Immune Responses *In Vivo*

There is a lack of suitable *in vivo* models for studies of the efficiency of genetic vaccines and the vast majority of studies are performed on the mouse model, in which it is sometimes difficult to predict whether the results obtained reliably predict similar vaccinations in humans because of the complexity of events occurring after genetic vaccination

A lack of suitable *in vivo* models has hampered studies of the efficiency of genetic vaccines in inducing antigen expression in human muscle cells and in inducing specific human immune responses. The vast majority of studies on the capacity of genetic vaccines to transfect muscle cells and to induce specific immune responses *in vivo* have employed a mouse model. Because of the complexity of events occurring after genetic vaccination, however, it is sometimes difficult to predict whether results obtained in the mouse model reliably predict the outcome of similar vaccinations in humans. The events required in successful genetic vaccination include transfection of the cells after delivery of the plasmid, expression of the desired antigen, processing the antigen in antigen presenting cells, presentation of the antigenic peptides in the context of MHC molecules, recognition of the peptide/MHC complex by T cell receptors, interactions of T cells with B cells and

professional antigen presenting cells and finally induction of specific T cell and B cell responses. All these events are likely to be somewhat differentially regulated in mouse and man.

The invention provides an *in vivo* model for human muscle cell transfection

Muscle tissue, obtained for example from cadavers, is transplanted subcutaneously into immunodeficient mice, which can be transplanted with tissues from other species without rejection. This model system is especially valuable because there is no *in vitro* culture system available for normal muscle cells. Muscle tissue, obtained for example from cadavers, is transplanted subcutaneously into immunodeficient mice. Immunodeficient mice can be transplanted with tissues from other species without rejection. Mice suitable for xenotransplantations include, but are not limited to, SCID mice, nude mice and mice rendered deficient in their genes encoding RAG1 or RAG2 genes. SCID mice and RAG deficient mice lack functional T and B cells, and therefore are severely immunocompromised and are unable to reject transplanted organs. Previous studies indicate that these mice can be transplanted with human tissues, such as skin, spleen, liver, thymus or bone, without rejection (Roncarolo et al. (1996) Semin. Immunol. 8: 207). After transplantation of human fetal lymphoid tissues into SCID mice, functional human immune system can be demonstrated in these mice, a model generally referred to as SCID-hu mice. When human muscle tissue is transplanted into SCID-hu mice, one can not only study transfection efficiency and expression of the desired antigen, but one can also study induction of specific human immune responses induced by genetic vaccines *in vivo*. In this case, muscle and lymphoid organs from the same donor are used. Fetal muscle also has an advantage in that it contains few mature lymphocytes of donor origin decreasing likelihood of graft versus host reaction.

Genetic vaccine vectors are introduced into the human muscle tissue to study the expression of the antigen of interest

Once the human muscle tissue is established in the mouse, genetic vaccine vectors are introduced into the human muscle tissue to study the expression of the antigen of interest. When studying transfection efficiency only, RAG deficient mice can be used. These mice never have mature B or T cells in the circulation, whereas "leakiness" of SCID phenotype has been demonstrated which may cause variation in the transplantation efficiency.

Model provides an efficient means to study gene expression in human muscle cells *in vivo*, despite the limited survival of the tissue in mice

The survival of human muscle tissue in mice is likely to be limited even in immuno-compromised mice. However, because expression studies can be performed within one or two days, this model provides an efficient means to study gene expression in human muscle cells *in vivo*. A modified SCID-hu mouse model with human muscle transplanted into these mice can be used to study human immune responses in mice *in vivo*.

Screening For Improved Delivery Of Vaccines

Identifying genetic vaccine vectors that are capable of being administered in a particular manner

For certain applications, it is desirable to identify genetic vaccine vectors that are capable of being administered in a particular manner, for example, orally or through the skin. The following screening methods provide suitable assays; additional assays are also described herein in conjunction with particular genetic vaccine properties for which the assays are especially suitable.

Screening for oral delivery either *in vitro* (based on Caco-2 cells) or *in vivo*

Screening for oral delivery can be performed either *in vitro* or *in vivo*. An example of an *in vitro* method is based on Caco-2 (human colon adenocarcinoma) cells which are grown in tissue culture. When grown on semipermeable filters, these cells spontaneously differentiate into cells that resemble human small intestine epithelium, both structurally and functionally. Genetic vaccine libraries and/or vectors can be placed on one side of the Caco-2 cell layer, and vectors that are able to move through the cell layer are detected on the opposite side of the layer.

Libraries can also be screened for amenability to oral delivery *in vivo*. For example, a library of vectors can be administered orally, after which target tissues are assayed for presence of vectors. Intestinal epithelium, liver, and the bloodstream are examples of tissues that can be tested for presence of library members. Vectors that are successful in reaching the target tissue can be recovered and, if further improvement is desired, used in succeeding rounds of reassembly (optionally in combination with other directed evolution methods described herein) and selection.

Apparatus which permits large numbers of vectors to be screened efficiently and can be used to study the effect of large numbers of agents *in vivo*

For screening a library of genetic vaccine vectors for ability to transfect cells upon injection into skin or muscle, the invention provides an apparatus which permits large numbers of vectors to be screened efficiently. This apparatus is based on 96-well format and is designed to transfer small volumes (2-5 μl) from a microtiter plate to skin or muscle of laboratory animals, such as mice and rats. Moreover, human muscle or skin transplanted into immunodeficient mice can be injected.

The apparatus is designed in such a way that the tips move to fit a microtiter plate. After the reagent of interest has been obtained from the plate, the distance of the tips from each other is decreased to 2-3 mm, enabling transfer of 96 reagents to an area of 1.6 cm x 2.4 cm to 2.4 cm x 3.6 cm. The volume of each sample transferred is electronically controlled. Each reagent is mixed with a marker agent or dye to enable recognition of injection site in the tissue. For example, gold particles of different sizes and shapes are mixed with the reagent of interest, and microscopy and immunohistochemistry can be used to identify each injection site and to study the reaction induced by each reagent. When muscle tissue is injected the injection site is first revealed by surgery.

This apparatus can be used to study the effects of large numbers of agents *in vivo*. For example, this apparatus can be used to screen efficiency of large numbers of different DNA vaccine vectors to transfect human skin or muscle cells transplanted into immunodeficient mice.

Enhanced Entry Of Genetic Vaccine Vectors Into Cells

Using stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly to efficiently improve the capacity of DNA to enter the cytoplasm and subsequently the nucleus of human cells

The methods involve subjecting to stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly polynucleotides which are involved in cell entry. Such polynucleotides are referred to herein as "transfer sequences" or "transfer modules." Transfer modules can be obtained which increase transfer in a cell-specific manner, or which act in a more general manner. Because the exact sequences that affect DNA binding and transfer are not often known, stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly may be the only efficient method to improve the capacity of DNA to enter the cytoplasm and subsequently the nucleus of human cells.

The stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly methods of the invention provide means for optimizing DNA sequences and the three-dimensional structure of the plasmids for ability to confer upon a vector the ability to enter a cell even in the absence of detailed information as to the mechanism by which this effect is achieved

The methods involve reassembling (&/or subjecting to one or more directed evolution methods described herein) at least first and second forms of a nucleic acid that comprises a transfer sequence. The first and second forms differ from each other in two or more nucleotides. Suitable substrates include, for example, transcription factor binding sites, CpG sequences, poly A, C, G, T oligonucleotides, non-stochastically generated nucleic acid building blocks ,and random DNA fragments such as, for example, genomic DNA, from human or other mammalian species. It has been suggested that cell surface proteins, such as the macrophage scavenger receptor, may act as receptors for specific DNA binding (Pisetsky (1996) Immunity 5: 303). It is not known whether these receptors recognize specific DNA sequences or whether they bind DNA in a sequence non-specific manner. However, GGGG tetrads have been shown to enhance DNA binding to cell surfaces (Id.). In addition to the DNA sequence, the three-dimensional structure of the plasmids may play a role in the capacity of these plasmids to enter cells. The stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly methods of the invention provide means for optimizing such sequences for ability to confer upon a vector the ability to enter a cell even in the absence of detailed information as to the mechanism by which this effect is achieved.

Clonal isolates of vectors bearing recombinant segments are used to infect separate cultures of cells and the percentage of vectors which enter cells is then determined by, for example, counting cells expressing a marker expressed by the vectors in the course of transfection

The resulting library of recombinant transfer modules are screened to identify at least one optimized recombinant transfer module that enhances the capability of a vector comprising the transfer module to enter a cell of interest. For example, vectors that include a recombinant transfer module can be contacted with a population of cells under conditions conducive to entry of the vector into the cells, after which the percentage of cells in the population which contain the nucleic acid vector is determined. In one aspect, the vector will contain a selectable or screenable marker to facilitate identification of cells which contain the

vector. In one aspect, clonal isolates of vectors bearing recombinant segments are used to infect separate cultures of cells. The percentage of vectors which enter cells can then be determined by, for example, counting cells expressing a marker expressed by the vectors in the course of transfection.

The reassembly (&/or one or more additional directed evolution methods described herein) and rescreening process can be repeated as necessary, until a transfer module that has sufficient ability to enhance transfer is obtained

Typically, the reassembly (&/or one or more additional directed evolution methods described herein) process is repeated by reassembling (&/or subjecting to one or more directed evolution methods described herein) at least one optimized transfer sequence with a further form of the transfer sequence to produce a further library of recombinant transfer modules. The further form can be the same or different from the first and second forms. The new library is screened to identify at least one further optimized recombinant vector module that exhibits an enhancement of the ability of a genetic vaccine vector that includes the optimized transfer module to enter a cell of interest.

The reassembly (&/or one or more additional directed evolution methods described herein) and rescreening process can be repeated as necessary, until a transfer module that has sufficient ability to enhance transfer is obtained. After one or more of reassembly (&/or one or more additional directed evolution methods described herein) and screening, vector modules are obtained which are capable of conferring upon a nucleic acid vector the ability to enter at least about 50 percent more target cells than a control vector which does not contain the optimized module, or at least about 75 percent more, or at least about 95 or 99 percent more target cells than a control vector.

For integration by homologous recombination, important factors are the degree and length of homology to chromosomal sequences, the frequency of such sequences in the genome, and the specific sequence mediating homologous recombination; for nonhomologous, illegitimate and site-specific recombination, recombination is mediated by specific sites on the therapy vector which interact with cell encoded recombination proteins

Although for vaccine purposes non-integrating vectors can be used, for some applications it may be desirable to use an integrating vector; for these applications DNA sequences that directly or indirectly affect the efficiency of integration can be included in the genetic vaccine vector. For integration by homologous recombination, important factors are the degree and length of homology to chromosomal sequences, as well as the frequency of such sequences in the genome (*e.g., Alu* repeats). The specific sequence mediating homologous recombination is also important, since integration occurs much more easily in transcriptionally active DNA. Methods and materials for constructing homologous targeting constructs are described by e.g., Mansour (1988) Nature 336:348; Bradley (1992) Bio/Technology 10:534. For nonhomologous, illegitimate and site-specific recombination, recombination is mediated by specific sites on the therapy vector which interact with cell encoded recombination proteins, e.g., Cre/Lox and Flp/Frt systems. See, e.g., Baubonis (1993) Nucleic Acids Res. 21:2025-2029, which reports that a vector including a LoxP site becomes integrated at a LoxP site in chromosomal DNA in the presence of Cre recombinase enzyme.

Optimization Of Genetic Vaccine Components

Optimizing properties that can influence the efficacy of a genetic vaccine in modulating an immune response in a mammalian system

Many factors can influence the efficacy of a genetic vaccine in modulating an immune response. The ability of the vector to enter a cell, for example, has a significant effect on the ability of the vector to modulate an immune response. The strength of an immune response is also mediated by the immunogenicity of an antigen expressed by a genetic vaccine vector and the level at which the antigen is expressed. The presence or absence of costimulatory molecules produced by the genetic vaccine vector can affect not only the strength, but also the type of immune response that arises due to introduction of the vector into a mammal. An increase in the persistence of a vector in an organism can lengthen the time of immunomodulation, and also makes feasible self-boosting vectors which do not

require multiple administrations to achieve long-lasting protection. The present invention provides methods for optimizing many of these properties, thus resulting in genetic vaccine vectors that exhibit improved ability to elicit the desired effect on a mammalian immune system.

The selection from large libraries using recursive cycles of reassembly (optionally in combination with other directed evolution methods described herein) to maximally access all the fortuitous but complex mechanisms that cannot be approached rationally

Genetic vaccines can contain a variety of functional components; desired sequences can be generated by (determined by) stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly, the empirical sequence evolution described in detail herein. The methods of the invention involve, in general, constructing a separate library for each of the major vector components by stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly of multiple homologous starting sequences, or other methods of generating a population of recombinants, resulting in a complex mixture of chimeric sequences. The best sequences are selected from these libraries using the high-throughput assays described below. After one or more cycles of selection from each of the single module libraries, the pools of the best sequences of different modules can be combined by stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly as long as the screens are compatible. The screens for promoter, enhancer, intron, transfer sequences, mammalian ori, bacterial ori and bacterial marker, and the like, can eventually be combined, resulting in co- optimization of the context of each sequence. An important aspect in these experiments is the selection from large libraries using recursive cycles of reassembly (optionally in combination with other directed evolution methods described herein) to maximally access all the fortuitous but complex mechanisms that cannot be approached rationally, such as DNA transfer into the cell.

A library of different vectors can be generated by assembling vector modules that provide promoters, cytokines, cytokine antagonists, chemokines, immunostimulatory sequences, and costimulatory molecules using assembly PCR and combinatorial molecular biology

Assembly PCR is a method for assembly of long DNA sequences, such as genes, non-stochastically generated nucleic acid building blocks, and fragments of plasmids. In contrast to PCR, there is no distinction between primers and template, because the non-stochastically

generated nucleic acid building blocks &/or fragments to be assembled prime each other. The library of vector modules obtained as described herein can be fused with promoters, which can themselves be optimized by the stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly methods of the invention. The resulting genes can be assembled combinatorially into DNA vaccine vectors, where each gene is expressed under a different promoter (e.g., a promoter derived from a library of experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) CMV promoters), and the vector library is screened as described herein to identify vectors which exhibit the desired effect on the immune system.

Properties that influence the efficacy or desirability of the vaccine

The methods of the invention are useful for obtaining genetic vaccines that are optimized for one or more of many properties that influence the efficacy or desirability of the vaccine. These properties include, but are not limited to, the following.

Episomal Vector Maintenance

Episomally replicating vectors are maintained in a cell for a longer period of time and permit the development of self-boosting vaccines

One property that one can optimize using the sequence reassembly methods of the invention is the ability of a genetic vaccine vector to replicate episomally in a mammalian cell. Episomal replication of a vaccine vector is advantageous in many situations. For example, episomally replicating vectors are maintained in a cell for a longer period of time than non-replicating vectors, thus resulting in an increased length of immune response modulation or increased delivery of a therapeutically useful protein. Episomal replication also permits the development of self-boosting vaccines which, unlike traditional vaccines, do not require multiple vaccine administrations. For example, a self-boosting vaccine vector can include an antigen-encoding gene which is under the control of an inducible control element which allows induction of antigen expression, and the corresponding immune response, in response to a specific stimulus. However, screening for naturally occurring vector modules which result in enhanced episomal maintenance using traditional approaches or attempts to rationally design mutants with improved properties would require many person-years of research. The invention provides methods for generating and screening orders of magnitude more diversity in a short time period.

Using stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly to recombine at least two forms of a nucleic acid which is capable of conferring upon a genetic vector the ability to replicate autonomously in mammalian cells

The ability of a genetic vaccine vector to replicate episomally can be optimized by using stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly to recombine at least two forms of a nucleic acid which is capable of conferring upon a genetic vector the ability to replicate autonomously in mammalian cells. The two or more forms of the episomal replication vector module differ from each other in two or more nucleotides. A library of recombinant episomal replication vector modules is produced, and the library is screened to identify one or more optimized replication vector modules which, when placed in a genetic vaccine vector, confer upon the vector an enhanced ability to replicate autonomously compared to a vector which contains a non-optimized episomal replication vector module.

Repetition of the stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly process at least once to identify modules which exhibit enhanced ability to confer episomal maintenance upon a vector containing the module

In one embodiment, the stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly process is repeated at least once using as a substrate an optimized episomal replication vector module obtained from a previous round of stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly. The optimized vector module obtained in the earlier round is reassembled (&/or subjected to one or more directed evolution methods described herein) with a further form of the vector module, which can be the same as one of the forms used in the earlier round, or can be a different form of a nucleic acid that functions as an episomal replication element. Again, a library of recombinant episomal replication vector modules is produced, and the screening process is repeated to identify those episomal replication modules which exhibit enhanced ability to confer episomal maintenance upon a vector containing the module.

Ability to replicate autonomously in eukaryotic cells- examples

Nucleic acids which are useful as substrates for the use of stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly to optimize episomal replication ability include any nucleic acid that is involved

in conferring upon a vector the ability to replicate autonomously in eukaryotic cells. For example, papillomavirus sequences E I and E2, simian virus 40 (SV40) origin of replication, and the like.

Genes from human papillomaviruses are exemplary episomal replication vector modules

Exemplary episomal replication vector modules that can be optimized using the methods of the invention are genes from human papillomaviruses (HPV) which are involved in episomal replication. HPV are non-tumorigenic viruses which replicate episomally in skin and are stably expressed *in vivo* for years. Bernard and Apt (1994) Arch. Dermatol. 130: 210.

Increased episomal maintenance of the HPV genes involved in episomal replication using directed evolution

Despite these *in vivo* properties, it has not been possible to maintain HPV episomally in tissue culture due to underreplication. The invention provides methods by which HPV genes involved in episomal maintenance can be optimized for use in genetic vaccine vectors. HPV genes involved in episomal replication include, for example, the El and E2 genes. Thus, according to one embodiment of the invention, either or both of the HPV E I and E2 genes are subjected to stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly to obtain a recombinant episomal replication module which, when placed in a nucleic acid vaccine vector, results in increased maintenance of the vector in mammalian cells. In one embodiment, the HPV E1 and E2 genes from different, but closely related, benign HPVs are used in a polynucleotide reassembly procedure, as shown, described &/or referenced herein (including incorporated by reference). For example, polynucletide shuffling of HPV El and E2 genes from closely related strains of HPV (such as, for example, HPV 2, 27, and 57) can be used to obtain a library of recombinant El and E2 genes which are then subjected to an appropriate screening method to identify those that exhibit improved episomal maintenance properties.

Identification, selection, enrichment of recombinant episomal replication vector modules that exhibit improved ability to mediate episomal maintenance

To identify recombinant episomal replication vector modules that exhibit improved ability to mediate episomal maintenance, members of the library of recombinant vector modules are inserted into vectors which are introduced into mammalian cells. The cells are propagated for at least several generations, after which cells that have maintained the vector

are identified. Identification can be accomplished, for example, employing a vector that includes a selectable marker. Cells containing the library members are propagated in the absence of selection for the selectable marker for at least several generations, after which selective pressure is added. Cells which survive selection are enriched for cells that harbor vectors which contain a recombinant vector module which enhances the ability of the vector to replicate episomally. DNA is recovered from the selected cells and introduced into bacterial host cells, allowing recovery of episomal, non- integrated vectors.

Screening by introducing to a vector containing a polynucleotide encoding an antigen that is present on the surface of the cell when expressed

In another embodiment of the invention, the screening step is accomplished by introducing members of the library of recombinant episomal replication vector modules into a vector that includes a polynucleotide that encodes an antigen which, when expressed, is present on the surface of a cell. The library of vectors is introduced into mammalian cells which are propagated for at least several generations, after which cells which display the cell surface antigen on the surface of the cell are identified. Such cells most likely harbor a genetic vaccine vector which enhances the ability of the vector to replicate autonomously.

Use of optimized recombinant episomal replication vector module to construct genetic vaccine vectors

Upon identifying cells which contain an episomally maintained vector, the optimized recombinant episomal replication vector module is obtained and used to construct genetic vaccine vectors. Cell surface antigens which are suitable for use in the screening methods are described above, and others are known to those of skill in the art. In one aspect, an antigen is used for which a convenient means of detection is available.

Exemplary cells for use in the screening methods

Cells which are suitable for use in the screening methods include both cultured mammalian cells and cells which are present in an animal. To screen for recombinant vector modules that are intended for use in humans, exemplary cells for screening purposes are human cells. Generally, initial screening is accomplished in cell culture, where processing of large libraries of experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) material is feasible. In one embodiment, cells which display a vector-encoded cell surface antigen on the cell surface are identified by flow

cytometry based cell sorting methods, such as fluorescence activated cell sorting. This approach allows very large numbers (> 10^7) cells to be evaluated in a single vial experiment.

Further testing for durability *in vivo* in an animal model

Constructs which replicate autonomously in cell culture and give rise to strong marker gene expression can be further tested for durability *in vivo* in an animal model. For example, mouse models for studies of human tissues in mice *in vivo* are described herein. Live pieces of human skin are xenotransplanted onto the back of SCID mice, allowing screening of the vector libraries for optimal properties in human cells *in vivo*. Recursive selection of episomal vectors will provide strong selection pressure for vectors that remain episomal, yet provide high level of gene expression.

Introducing a genetic vaccine vector into a mammal that has a functional human immune system and testing for the existence of an immune response against the antigen

In another embodiment, the screening step involves introducing a genetic vaccine vector which includes the recombinant episomal replication vector module, as well as polynucleotide that encodes an antigen or pharmaceutically useful protein, into a mammal that has a functional human immune system. The animal is then tested for the existence of an immune response against the antigen. In one embodiment, the mammals used for such assays are non-human mammals that have a functional human immune system. For example, a functional human immune system can be created in an immunodeficient mouse by introducing one or more of a human fetal tissue selected from the group consisting of liver, thymus, and bone marrow (Roncarolo et al. (1996) Semin. Immunol. 8: 207).

Episomally maintained vectors result in high signal-to-noise ratios upon FACS selection and significantly improve the possibility to recover the plasmids from a small number of selected cells

Stable episomal vectors which are obtained using the methods of the invention are useful not only as genetic vaccines, but also are useful tools in other library screening applications. In contrast to randomly integrating and transient vectors, episomally maintained vectors result in high signal-to-noise ratios upon FACS selection, and they also significantly improve the possibility to recover the plasmids from a small number of selected cells.

Evolution Of Optimized Promoters For Expression Of An Antigen

Optimizing the promoter and/or other control sequence to improve the efficacy of genetic vaccinations, reduce the amount of DNA required for protective immunity and thereby the

cost of vaccination, control the type of cell in which the gene is expressed, and/or the timing of the antigen expression

In another embodiment, the invention provides methods of optimizing vector modules such as promoters and other gene expression control signals. Usually, a coding sequence for an antigen that is delivered by a genetic vaccine is operably linked to an additional sequence, such as a regulatory sequence, to ensure its expression. These regulatory sequences can include one or more of the following: an enhancer, a promoter, a signal peptide sequence, an intron and/or a polyadenylation sequence. A desirable goal is to increase the level of expression of functional expression product relative to that achieved with conventional vectors. The efficacy of a genetic vaccine vector often depends on the level of expression of an antigen by the vaccine vector. An optimized promoter and/or other control sequence is likely to result in improved efficacy of genetic vaccinations, reduce the amount of DNA required for protective immunity and thereby the cost of vaccination.

Moreover, it is sometimes desirable to have control over the type of cell in which a gene is expressed, and/or the timing of antigen expression. The methods of the invention provide for optimization of these and other factors which are influenced by promoters and other control sequences.

Improving expression by increasing the rate of production of an expression product, decreasing the rate of degradation of the expression product, or improving the capacity of expression product to perform its intended function using stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly of polynucleotides involved in control of gene expression

Improved expression of selection markers can be achieved by performing stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly, for example. Expression can effectively be improved by a variety of means, including increasing the rate of production of an expression product, decreasing the rate of degradation of the expression product or improving the capacity of the expression product to perform its intended function. The methods involve subjecting to stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly polynucleotides which are involved in control of gene expression. At least first and second forms of a nucleic acid that comprises a control sequence, which forms differ from each other in two or more nucleotides, are reassembled (&/or subjected to one or more

directed evolution methods described herein) as described above. The resulting library of recombinant transfer modules are screened to identify at least one optimized recombinant control sequence that exhibits enhanced strength, inducibility, or specificity.

Introduction of the recombinant segments at the level of fragments (non-tochastically generated &/or randomly generated) and *in vitro*

The substrates for reassembly (&/or one or more additional directed evolution methods described herein) can be the full-length vectors, or fragments thereof, which include a coding sequence and/or regulatory sequences to which the coding sequence is operably linked. The substrates can include variants of any of the regulatory and/or coding sequence(s) present in the vector. If reassembly (&/or one or more additional directed evolution methods described herein) is effected at the level of fragments, the recombinant segments should be reinserted into vectors before screening. If reassembly (&/or one or more additional directed evolution methods described herein) proceeds *in vitro*, vectors containing the recombinant segments are usually introduced into cells before screening. An example of a vector suitable for use in screening of experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) promoters and other regulatory regions is shown, described &/or referenced herein (including incorporated by reference).

Using an easily detected selection marker (green fluorescent protein, cell surface protein) when an additional or substitute marker is required

Cells containing the recombinant segments can be screened by detecting expression of the gene encoded by the selection marker. For purposes of selection and/or screening, a gene product expressed from a vector is sometimes an easily detected marker rather than a product having an actual therapeutic purpose, e.g., a green fluorescent protein (see, Crameri (1996) Nature Biotechnol. 14: 315-319) or a cell surface protein. For example, if this marker is green fluorescent protein, cells with the highest expression levels can be identified by flow cytometry-based cell sorting. If the marker is a cell surface protein, the cells are stained with a reagent having affinity for the protein, such as antibody, and again analyzed by flow cytometry-based cell sorting. However, some genes having a therapeutic purpose, e.g., drug resistance genes, themselves provide a selectable marker, and no additional or substitute marker is required. Alternatively, the gene product can be a fusion protein comprising any combination of detection and selection markers. Internal reference marker genes can be

included on the vector to detect and compensate for variations in copy number or insertion site.

Further round of reassembly (&/or one or more additional directed evolution methods described herein) and screening.

Recombinant segments from the cells showing highest expression of the marker gene can be used as some or all of the substrates in a further round of reassembly (&/or one or more additional directed evolution methods described herein) and screening, if additional improvement is desired.

Constitutive Promoters

Evolving control sequences (promoters, enhancers, etc.) to express a gene of interest at a higher level than is a gene operably linked to a non-evolved control sequences

The invention provides methods of evolving nucleotide sequences that are capable of directing constitutive expression of a gene of interest which is operably linked to the control sequence. Typically, the control sequences, which can include promoters, enhancers, and the like, are evolved so that a gene of interest is expressed at a higher level than is a gene operably linked to a non-evolved control sequence. To screen for control sequences which are of increased strength, a recombinant library of control sequences can be introduced into a population of cells and the level of expression of a detectable marker operably linked to the control sequences determined. In one aspect, the optimized promoter is capable of expressing an operably linked gene at a level that is at least about 30% greater than that of a control promoter construct, or the optimized promoter is at least about 50% stronger than a control, or at least about 75% or more stronger than a control promoter.

Using improved CMV promoter/enhancer elements (SV40 and Sra) to express foreign genes both in animal models and in clinical applications

Examples of promoters which can be used as substrates in the methods include any constitutive promoter that functions in the intended host cell. The major immediate-early (IE) region transcriptional regulatory elements, including promoter and enhancer sequences (the promoter/enhancer region), of cytomegalovirus (CMV) is widely used for regulating transcription in vectors used for gene therapy because it is highly active in a broad range of cell types. Optimized CMV transcriptional regulatory elements which direct increased levels of antigen expression is generated by the recursive reassembly (&/or one or more additional directed evolution methods described herein) methods of the invention, resulting in improved

efficacy of gene therapy. As the CMV promoter and enhancer is active in human and animal cells, the improved CMV promoter/enhancer elements are used to express foreign genes both in animal models and in clinical applications. Other constitutive promoters that are amenable to use in the claimed methods include, for example, promoters from SV40 and SRα, and other promoters known to those of skill in the art.

Creating a library of chimeric transcriptional regulatory elements through stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly of wild-type sequences from two or more of the five related strains of CMV, obtaining the promoter, enhancer and first intron sequences of the IE region through PCR of the CMV strains

In one embodiment, a library of chimeric transcriptional regulatory elements is created by stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly of wild-type sequences from two or more of the five related strains of CMV. The promoter, enhancer and first intron sequences of the IE region are obtained by PCR from the CMV strains: human VR-53 8 strain AD169 (Rowe (195 6) Proc. Soc. Exp. Biol. Med. 92:418; human V-977 strain Towne (Plotkin (1975) Infect. Immunol. 12:521-527); rhesus VR-677 strain 68-1 (Asher (1969) BacterioL Proc. 269:91); vervet VR-706 strain CSG (Black (1963) Proc. Soc. Exp. Biol. Med. 112:60 1); and, squirrel monkey VR-1398 strain SqSHV (Rangan (1980) Lab. Animal Sci. 30:532). The promoter/enhancer sequences of the human CMV strains are 95% homologous, and share 70% homology with the sequences of the monkey isolates, allowing the use of polynucleotide reassembly (optionally in combination with other directed evolution methods described herein) to generate a library great diversity. Following reassembly (optionally in combination with other directed evolution methods described herein), the library is cloned into a plasmid backbone and used to direct transcription of a marker gene in mammalian cells. An internal marker under the control of a native promoter is typically included in the plasmid vector, which will allow analysis and sorting of cells harboring equal numbers of vectors.

Expression markers, such as green fluorescent protein (GFP) and CD86 (also known as B7.2, see Freeman (1993) J Exp. Med 178:2185, Chen (1994) J Immunol. 152:4929) can also be used. In addition, transfection of SV40 T antigen-transformed cells can be used to amplify a vector which contains an SV40 origin of replication. The transfected cells are screened by FACS sorting to identify those which express high levels of the marker gene,

normalized against the internal marker to account for differences in vector copy numbers per cell. If desired, vectors carrying optimal, recursively reassembled (&/or subjected to one or more directed evolution methods described herein) promoter sequences are recovered and subjected to further cycles of reassembly (optionally in combination with other directed evolution methods described herein) and selection.

Cell-Specific Promoters

Reducing the risk of autoimmune disorder following introduction of foreign antigens into host cells and providing for efficient induction of protective immunity through the expression of genetic vaccines in professional APCs, such as dendritic cells and macrophages

One of the safety concerns associated with genetic vaccines has been the possibility of autoimmune disorders following introduction of foreign antigens into host cells. This risk can be reduced if the pathogen antigen is specifically expressed in professional APCs that express the proper costimulatory molecules. Although it is somewhat debatable which cells are the most important cells expressing the pathogen antigen following genetic vaccinations, it is likely that professional APCs are involved. It has been shown that blood monocytes express antigen following intramuscular injection of genetic vaccine vectors, and dendritic cells derived from lymph nodes of vaccinated animals efficiently induced antigen-specific T cell activation (C. Bona, The First Gordon Conference on Genetic Vaccines, Plymouth, NH, July 21, 1997). These data, together with previous studies indicating that small number of dendritic cells expressing antigen or antigenic peptides is sufficient to induce activation of antigen-specific T cells (Thomas and Lipsky, Stem Cells 14: 196, 1996), support the conclusion that genetic vaccines specifically expressed in professional APC, such as dendritic cells and macrophages, are likely to provide efficient induction of protective immunity with minimized chance of adverse effects.

Methods for obtaining promoters and enhancers that induce high expression levels specifically in professional APCs, exploiting natural diversity as a source of substrates for stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly

The present invention provides methods of obtaining promoters and enhancers that induce high expression levels specifically in professional APCs. Previously existing APC-specific vectors did not provide sufficient expression levels following genetic vaccinations. The methods involve performing stochastic (e.g. polynucleotide shuffling & interrupted

synthesis) and non-stochastic polynucleotide reassembly as described above using as substrates different forms of a nucleic acid that comprises an APC- specific promoter or other control signal. Suitable promoters include, for example, the MHC Class II, and the CD11b, CD11c, and CD40 promoters. Natural diversity of the promoters can be exploited as a highly appropriate source of substrates for the stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly. For example, genomic DNA from monkeys, pigs, dogs, cows, cats, rabbits, rats and mice, can be obtained, and the proper sequences obtained by using multiple PCR primers specific for the most conserved regions based on known sequence information. The selection of the optimal promoters can be done in monocytic or B cell lines, such as U937, HL60 or Jijoye, using FACS- sorting. In addition, $SV40^+$ cell lines, such as COS-1 and COS-7, can be used to improve the recovery of the plasmids. Further analysis can be undertaken in human dendritic cells obtained by culturing peripheral blood monocytes in the presence of IL-4 and GM-CSF as described (Chapuis et al. (1997) Eur. J Immunol. 27: 43 1).

Inducible Promoters

Using stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly of two substrates, such as tetracycline and hormone inducible expression systems, to increase the expression level and inducibility *in vivo* of the promoter controlling transgene expression

A particularly desirable property of a genetic vaccines would be an ability to induce the promoter controlling transgene expression simply by taking an innocuous oral drug, resulting in a boost of the immune response. Essential requirements for inducible promoters are low base-line expression and strong inducibility. Several promoters with exquisite *in vitro* regulation exist, but the expression level and inducibility of each is too low to be useable *in vivo*. The invention overcome these problems by stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly using as substrates two or more variants of a nucleic acid that functions as an inducible control sequence. Suitable substrates include, for example, tetracycline and hormone inducible expression systems, and the like. Hormones that have been used to regulate gene expression include, for example, estrogen, tomoxifen, toremifen and ecdysone (Ramkumar and Adler (1995) Endocrinology 136: 536-542). Libraries of recombinant inducible promoters are screened as described above in the presence and absence of the inducer.

Tetracycline responsive system provides possibilities to induce and turn off gene expression (ecdysone responsive element another candidate)

The most commonly used inducible gene expression protocol is the tetracycline responsive system, which provides possibilities to both induce and turn off gene expression (Gossen and Bujard (1992) Proc. Nat'l. Acad. Sci. USA 89: 5547; Gossen et al. (1995) Science 268: 1766). A repressor gene is located on the plasmid and binds to an operator in the promoter. Tetracycline or doxycycline modulates the binding ability of the repressor. Interestingly, four amino acid changes convert the repressor into an activator. In addition to the tetracycline responsive system, other candidates for inducible promoter evolution include the ecdysone responsive element (No et al., Proc. Nat'l. Acad Sci. USA 93:3346,1997).

Inducible promoters provide a means by which a vaccine dose can be administered subsequent to the initial administration simply by ingestion of a reagent that causes induction of the inducible promoter

Inducible promoters such as those obtained using the methods of the invention are useful in autoboost vaccines. Particularly when combined with a stably maintained episomal vector obtained as described above, the inducible promoters provide a means by which a vaccine dose can be administered subsequent to the initial administration simply by ingestion of a reagent that causes induction of the inducible promoter. A flow cytometry-based screening protocol that is suitable for optimization of inducible promoters is diagrammed herein.

Testing the functionality of autoboosting vaccines in a mouse model

The functionality of autoboosting vaccines can be tested in a mouse model such as that described above. Genetic vaccine vectors are injected into the skin of normal mice and into human skin in SCID-human skin mice. A gene encoding hepatitis B surface antigen (HBsAg) or other surface antigen is incorporated into these vectors enabling direct measurements of the levels of antigen produced, because HBsAg levels can be measured in cell culture supernates and in the circulation of the mice. The drug inducing the expression of the antigen is given after 1, 2, 4 and 6 weeks, and the expression levels of HBsAg are studied. Moreover, the levels of anti-HBsAg antibodies are measured. The mice are also injected with a vector containing a pathogen antigen discovered by ELI, and specific immune responses are followed.

In vivo assessment of functionality of autoboosting genetic vaccines in human immune system using SCID-human skin model with SCID-hu mouse model

Combining the SCID-human skin model with traditional SCID-hu mouse model (Roncarolo et al., Semin. Immunol. 8: 207, 1996) allows the assessment of functionality of autoboosting genetic vaccines in human immune system *in vivo*, and also allows measurements of human Ab responses *in vivo*. This model can also be used to assess production of HBsAg after oral boosting of novel genetic vaccine vectors harboring the gene encoding HBsAg.

Evolution Of Binding Polypeptides That Enhance Specificity And Efficiency Of Genetic Vaccines

The present invention also provides methods for obtaining recombinant nucleic acids that encode polypeptides which can enhance the ability of genetic vaccines to enter target cells. Although the mechanisms involved in DNA uptake are not well understood, the methods of the invention enable one to obtain genetic vaccines that exhibit enhanced entry to cells, and to appropriate cellular compartments.

Enhancing the efficiency and specificity of a genetic vaccine nucleic acid uptake by a given cell type by coating the nucleic acid with an evolved protein that binds to the genetic vaccine nucleic acid, and is also capable of binding to the target cell

In one embodiment, the invention provides methods of enhancing the efficiency and specificity of a genetic vaccine nucleic acid uptake by a given cell type by coating the nucleic acid with an evolved protein that binds to the genetic vaccine nucleic acid, and is also capable of binding to the target cell. The vector can be contacted with the protein *in vitro* or *in vivo*. In the latter situation, the protein is expressed in cells containing the vector, optionally from a coding sequence within the vector. The nucleic acid binding proteins to be evolved usually have nucleic acid binding activity but do not necessarily have any known capacity to enhance or alter nucleic acid DNA uptake.

DNA binding proteins that can be used in these methods

DNA binding proteins which can be used in these methods include, but are not limited to, transcriptional regulators, enzymes involved in DNA replication (*e.g., recA*) and reassembly (&/or one or more additional directed evolution methods described herein), and proteins that serve structural functions on DNA (e.g., histones, protamines). Other DNA binding proteins that can be used include the phage 434 repressor, the lambda phage cI and

cro repressors, the E. coli CAP protein, myc, proteins with leucine zippers and DNA binding basic domains such as fos and jun; proteins with'POU' domains such as the Drosophila paired protein; proteins with domains whose structures depend on metal ion chelation such as Cys_2His_2 zinc fingers found in TFIIIA, $Zn_2(Cys)_6$ clusters such as those found in yeast *Gal*4, the Cys_3 His box found in retroviral nucleocapsid proteins, and the $Zn_2(Cys)_8$ clusters found in nuclear hormone receptor-type proteins; the phage P22 Arc and Mnt repressors (see Knight et al. (1989) J Biol. Chem. 264: 3639- 3642 and Bowie & Sauerkl 989) J Biol. Chem. 264: 7596-7602. RNA binding proteins are reviewed by Burd & Dreyfuss (1994) Science 265: 615-62 1, and include HIV Tat and Rev.

Formats for performing reassembly (&/or one or more additional directed evolution methods described herein)

As in other methods of the invention, evolution of DNA binding proteins toward acquisition of improved or altered uptake efficiency is effective by one or more cycles of reassembly (&/or one or more additional directed evolution methods described herein) and screening. The starting substrates can be nucleic acid segments encoding natural or induced variants of one or nucleic acid binding proteins, such as those mentioned above. The nucleic acid segments can be present in vectors or in isolated form for the reassembly (&/or one or more additional directed evolution methods described herein) step. reassembly (&/or one or more additional directed evolution methods described herein) can proceed through any of the formats described herein.

For screening purposes, the reassembled (&/or subjected to one or more directed evolution methods described herein) nucleic acid segments are typically inserted into a vector, if not already present in such a vector during the reassembly (&/or one or more additional directed evolution methods described herein) step.

Including binding site in vector for DNA binding protein recognizing a specific binding site

The vector generally encodes a selective marker capable of being expressed in the cell type for which uptake is desired. If the DNA binding protein being evolved recognizes a specific binding site (*e.g.*, *lac*I binding protein recognizes *lac*O), this binding site can be included in the vector. Optionally, the vector can contain multiple binding sites in tandem.

Transforming vectors containing recombinant segments into host cells and lysing cells under mild conditions that do not disrupt binding of vectors to DNA binding proteins

The vectors containing different recombinant segments are transformed into host cells, usually E. coli, to allow recombinant proteins to be expressed and bind to the vector encoding their genetic material. Most cells take up only a single vector and so transformation results in a population of cells, most of which contain a single species of vector. After an appropriate period to allow for expression and binding, cells are lysed under mild conditions that do not disrupt binding of vectors to DNA binding proteins. For example, a lysis buffer of 35mM HEPES (pH 7.5 with KOH), 0.1mM EDTA, 100mM Na glutamate, 5% glycerol, 0.3mg/ml BSA, 1mM DTT, and 0.1mM PMSF) plus lysozyme (0.3-ml at 10 mg/ml) is suitable (see Schatz et al., US 5,338,665). The complexes of vector and nucleic acid binding protein are then contacted with cells of the type for which improved or altered uptake is desired under conditions favoring uptake. Suitable recipient cells include the human cell types that are common targets in DNA vaccination. These cells include muscle cells, monocytes/macrophages, dendritic cells, B cells, Langerhans cells, keratinocytes, and the M-cells of the gut. Cells from mammals including, for example, human, mouse, and monkey can be used for screening. Both primary cells and cells obtained from cell lines are suitable.

Recovery of cells expressing marker and enriching for recombinant segments for further rounds of selection

After incubation, cells are plated with selection for expression of the selective marker present in the vector containing the recombinant segments. Cells expressing the marker are recovered. These cells are enriched for recombinant segments encoding nucleic acid binding proteins that enhance uptake of vectors encoding the respective recombinant segments. The recombinant segments from cells expressing the marker can then be subjected to a further round of selection. Usually, the recombinant segments are first recovered from cells, e.g., by PCR amplification or by recovery of the entire vectors. The recombinant segments can then be reassembled (&/or subjected to one or more directed evolution methods described herein) with each other or with other sources of DNA binding protein variants to generate further recombinant segments. The further recombinant segments are screened in the same manner as before.

Using stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly to evolve, particularly, the carboxy- and amino-terminal peptide extensions of the histone protein, to increase the efficiency of DNA transfer into the cells

One example of a method to evolve an optimized nucleic acid binding domain involves the reassembly (optionally in combination with other directed evolution methods described herein) of histone genes. Histone-condensed DNA can result in increased gene transfer into cells. See, e.g., Fritz et al. (1996) Human Gene Therapy 7: 1395-1404. Thus, stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly can be used to evolve the histone protein, particularly the carboxy- and amino-terminal peptide extensions, to increase the efficiency of DNA transfer into cells. In this approach, the histone is encoded by the DNA to which it will be bound.

Construction of the histone library

The histone library can be constructed by, for example, 1) reassembly (optionally in combination with other directed evolution methods described herein) of many related histone genes from natural diversity, 2) addition of random or partially randomized peptide sequences at the N- and C-terminal sequences of the histone, 3) by addition of pre-selected protein-encoding regions to the N- or C-termini, such as whole cDNA libraries, nuclear protein ligand libraries, etc. These proteins can be partially randomized and linked to the histone by a library of linkers.

Starting substrates for evolving nucleic acid binding sites contain variant binding sites and recombinant forms of these sites are screened as a component of a vector that also encodes a nucleic acid binding protein

In a variation of the above procedure, a binding site recognized by a nucleic acid binding protein can be evolved instead of, or as well as, the nucleic acid binding protein. Nucleic acid binding sites are evolved by an analogous procedure to nucleic acid binding proteins except that the starting substrates contain variant binding sites and recombinant forms of these sites are screened as a component of a vector that also encodes a nucleic acid binding protein.

When the evolved DNA binding protein does not have a high degree of sequence specificity and it is unknown precisely which sites of the vector used in screening are bound by the protein, the vector should include all or most of the screening vector sequences together with additional sequences required to effect vaccination or therapy

Evolved nucleic acid segments encoding DNA binding proteins and/or evolved DNA binding sites can be included in genetic vaccine vectors. If the affinity of the DNA binding protein is specific to a known DNA binding site, it is sufficient to include that binding site

and the sequence encoding the DNA binding protein in the genetic vaccine vector together with such other coding and regulatory sequences are required to effect gene therapy. In some instances, the evolved DNA binding protein may not have a high degree of sequence specificity and it may be unknown precisely which sites on the vector used in screening are bound by the protein. In these circumstances, the vector should include all or most of the screening vector sequences together with additional sequences required to effect vaccination or therapy. An exemplary selection scheme which employs M 13 protein VIII is shown, described &/or referenced herein (including incorporated by reference).

Target cells of interest

Target cells of interest include, for example, muscle cells, monocytes, dendritic cells, B cells, Langerhans cells, keratinocytes, M-cells of the gut, and the like. Cell- specific ligands that are suitable for use with each of the cell types are known to those of skill in the art. For example, suitable proteins to direct binding to antigen presenting cells include CD2, CD28, CTLA-4, CD40 ligand, fibrinogen, factor X, ICAM-1, β-glycan (zymosan), and the Fc portion of immunoglobulin G. (Weir's Handbook of Experimental Immunology, Eds. L.A. Herzenberg, D.M. Weir, L.A. Herzenberg, C. Blackwell, 5th edition, volume IV, chapters 156 and 174) because their respective ligands are present on APCs, including dendritic cells, monocytes/macrophages, B cells, and Langerhans cells. Bacterial enterotoxins or subunits thereof are also of interest for targeting purposes.

LPS facilitates the interaction between vector and monocytes and is also likely to act as an adjuvant, further potentiating the immune responses

The ability of the vectors to enter and activate APC, such as monocytes, can also be enhanced by coating the vectors with small quantities of lipopolysaccharide (LPS). This facilitates the interaction between vector and monocytes, which have a cell surface receptor for LPS. Due to its immunostimulatory activities, LPS is also likely to act as an adjuvant, thereby further potentiating the immune responses.

Receptor binding components of enterotoxins can be evolved for improved attachment to cell surface receptors, improved entry to and transport across the cells of the intestinal epithelium, and improved binding to, and activation of, B cells or other APCs

Enterotoxins produced by certain pathogenic bacteria are useful as agents that bind cells and thus enhance delivery of vaccines, antigens, gene therapy vectors and pharmaceutical proteins. In an exemplary embodiment of the invention, receptor binding

components of enterotoxins derived from Vibrio cholerae and enterotoxigenic strains of E. coli are evolved for improved attachment to cell surface receptors and for improved entry to and transport across the cells of the intestinal epithelium. In addition, they can be evolved for improved binding to, and activation of, B cells or other APCs. An antigen of interest can be fused to these toxin subunits to illustrate the feasibility of the approach in oral delivery of proteins and to facilitate the screening of evolved enterotoxin subunits. Examples of such antigens include growth hormone, insulin, myelin basic protein, collagen and viral envelope proteins.

Vectors that contain the library of recombinant enterotoxin binding moiety nucleic acids are transfected into a population of host cells, wherein the recombinant enterotoxin binding moiety nucleic acids are expressed to form recombinant enterotoxin binding moiety polypeptides

These methods involve reassembling (&/or subjecting to one or more directed evolution methods described herein) at least first and second forms of a nucleic acid which comprises a polynucleotide that encodes a receptor binding moiety, e.g., a non-toxic receptor binding moiety, of an enterotoxin. The first and second forms differ from each other in two or more nucleotides, so the stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly results in production of a library of recombinant enterotoxin binding moiety nucleic acids. Suitable enterotoxins include, for example, a V cholerae enterotoxin, enterotoxins from enterotoxigenic strains of E. coli, salmonella toxin, shigella toxin and campylobacter toxin. Vectors that contain the library of recombinant enterotoxin binding moiety nucleic acids are transfected into a population of host cells, wherein the recombinant enterotoxin binding moiety nucleic acids are expressed to form recombinant enterotoxin binding moiety polypeptides. In one embodiment, the recombinant enterotoxin binding moiety polypeptides are expressed as fusion proteins on the surface of bacteriophage particles. The recombinant enterotoxin binding moiety polypeptides can be screened by contacting the library with a cell surface receptor of a target cell and determining which recombinant enterotoxin binding moiety polypeptides exhibit enhanced ability to bind to the target cell receptor. The cell surface receptor can be present on the surface of a target cell itself, or can be attached to a different cell, or binding can be tested using cell surface receptor that is not associated with a cell. Examples of suitable cell surface receptors include, for example, Gm I. Similarly, one can evolve bacterial superantigens for

altered (increased or decreased) binding to T cell receptor and MHC class H molecules. These superantigens activate T cells in an antigen nonspecific manner.

Superantigens binding to T cell receptor/MHC class II molecules include Staphylococcal enterotoxin B, *Urtica dioica* superantigen (Musette et al. (1996) Eur. J 1mmunol. 26:618- 22) and Staphylococcal enterotoxin A (Bavari et al. (1996) J Infect. Dis. 174:338-45). Phage display has been shown to be effective when selecting superantigens that bind MHC class H molecules (Wung and Gascoigne (1997) J 1mmunol. Methods. 204:33- 41).

Both CT and CT-B have been shown to have potent adjuvant activities *in vivo* and they enhance immune responses after oral delivery of antigens and vaccines

Cholera toxin (CT) is an oligomeric protein of 84,000 daltons which consists of one toxic A subunit (CT-A) covalently linked to five B subunits (CT- B). CT-B functions as the-receptor binding component and binds to G_{M1}, ganglioside receptors on mammalian cell surfaces. The toxic A-subunit is not necessary for the function of CT, and in the absence of CT-A, functional CT-B pentamers can form (Lebens and Holingren (1994) Dev. Biol. Stand. 82: 215-227). Both CT and CT-B have been shown to have potent adjuvant activities *in vivo* and they enhance immune responses after oral delivery of antigens and vaccines (Czerkinsky et al. (1996) Ann. NY Acad. Sci. 778: 185-93; Van Cott et al. (1996) Vaccine 14: 392-8). Moreover, a single dose of CT-B conjugated to myelin basic protein prevented onset of autoimmune encephalomyelitis (EAE), a murine model of multiple sclerosis (Czerkinsky et al., supra.). Furthermore, feeding animals with myelin basic protein conjugated to CT-B after the onset of clinical symptoms (7 days) attenuated the symptoms in these animals. Other bacterial toxins, such as enterotoxins of E. coli, Salmonella toxin, Shigella toxin and Campylobacter toxin, have structural similarities with CT. Enterotoxins of E. coli have the same A-B structure as CT and they also have sequence homology and share functional similarities.

Family stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly is feasible among enterotoxin- encoding nucleic acids from different bacterial species

Bacterial enterotoxins can be evolved for improved affinity and entry to cells by polynucleotide (e.g. gene, promoter, enhancer, intron, & the like) reassembly (optionally in combination with other directed evolution methods described herein). The similarity of E.

coli-derived enterotoxin subunit and CT-B is 78%, and several completely conserved regions of more than eight nucleotides can be found. B subunits from two different strains of E. coli are 98% homologous both at sequence and protein levels. Thus, family stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly is feasible among enterotoxin- encoding nucleic acids from different bacterial species.

Screen the secretion of chimeric proteins by V cholerae by culturing the bacteria in agar in the presence of monoclonal antibodies specific for the antigen that was fused to the toxins

The libraries of experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) toxin subunits can be expressed in a suitable host cell, such as V cholerae. For safety reasons, strains in which the toxic CT-A is deleted can be used. An antigen of interest can be fused to the receptor-binding subunit. Secretion of chimeric proteins by V cholerae can be screened by culturing the bacteria in agar in the presence of monoclonal antibodies specific for the antigen that was fused to the toxins and the level of secretion is detected as immunoprecipitation in the agar around the colonies.

Evolving for improved binding to the G_{M1}, ganglioside receptor and other receptors, detecting binding between receptor and chimeric fusion proteins with a monoclonal antibody specific for the antigen that was fused to the toxin

One can also add G_{M1}, ganglioside receptors to the agar in order to detect colonies secreting functional enterotoxin subunits. Colonies producing significant levels of the fusion protein are then cultured in 96-well plates, and the culture medium is tested for the presence of molecules capable of binding to cells or receptors in solution. Binding of chimeric fusion proteins to G_{M1}, ganglioside receptors on cell surface or in solution can be detected by a monoclonal antibody specific for the antigen that was fused to the toxin. The assay using whole cells has the advantage that one may evolve for improved binding also to receptors other than the G_{M1}, ganglioside receptor. When increasing concentrations of wild-type enterotoxins are added to these assays, one can detect mutants that bind to receptors with improved affinities. Affinity and specificity of toxin binding can also be determined by surface plasmon resonance (Kuziemko et al. (1996) Biochemistry 35: 6375- 84).

Advantage of large scale production and avoidance of problems associated with expression on phage in the bacterial expression system

The advantage of the bacterial expression system is that the fusion protein is secreted by bacteria that could potentially be used in large scale production. Moreover, because the fusion protein is in solution during selection, possible problems associated with expression on phage (such as bias towards selection of mutants that only function on phage) can be avoided.

In phage display, mutants can be easily further selected in *in vivo* assays when screening to identify enterotoxins with improved affinities

Nevertheless, phage display is useful for screening to identify enterotoxins with improved affinities. A library of experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) mutants can be expressed on phage, such as M 13, and mutants with improved affinity are selected based on binding to, for example, G_{M1} ganglioside receptors in solution or on a cell surface. The advantage of this approach is that the mutants can be easily further selected in *in vivo* assays as discussed below. A screening approach using fusion to M 13 protein VIII is diagrammed herein.

The recombinant binding moiety is expressed in the cells and binds to the nucleic acid binding domain to form a vector-binding moiety complex

Finally, the resulting evolved enterotoxin can be fused with DNA binding protein, and genetic vaccine vectors are coated with this fusion protein. The stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly can be done either separately, in which case the two domains are assembled after reassembly (optionally in combination with other directed evolution methods described herein), or in a combined reaction. reassembly (optionally in combination with other directed evolution methods described herein) results in production of a library of recombinant binding moiety nucleic acids which can be screened by transfecting vectors which contain the library, as well as a binding site specific for the nucleic acid binding domain, into a population of host cells. The binding moiety is expressed in the cells and binds to the nucleic acid binding domain to form a vector-binding moiety complex. Host cells can then be lysed under conditions that do not disrupt binding of the vector- binding moiety complex.

Optimized recombinant binding moiety nucleic acids are isolated from cells containing the vector

The vector -binding moiety complex can then be contacted with a cell of interest, after which cells are identified that contain a vector and the optimized recombinant binding moiety nucleic acids are isolated from the cells.

Increasing the number of copies of target DNA taken into those cells that initially take up the same DNA (mammalian cells)

Another method for obtaining enhanced uptake of a target DNA by mammalian cells is also provided by the invention. Specifically, the method increases the number of copies of target DNA taken into those cells that initially take up the same DNA.

Cells that take up the target molecule of DNA (cell surface expression of membrane-associated DNA binding domains) will express the factor and have increased specific affinity for target DNA that remains extracellular, while cells that did not take up DNA will be at a competitive disadvantage as they will not bear the cell surface target DNA- specific binding domain, which is required for specifically mediated DNA uptake

The method uses cell surface expression of membrane-associated DNA binding domains of, for example, transcription factors, that are encoded in the target DNA sequence, which also includes the cognate recognition sequence for the binding domain. Uptake of one molecule of target DNA into a cell (by any process, passive uptake, electroporation, osmotic shock, other stress) will lead to transcription of the gene encoding the polynucleotide binding domain. The gene encoding the binding domain is engineered so that the binding domain is expressed in a membrane anchored form. For example, a hydrophobic stretch of amino acids can be encoded at the carboxyl terminus of the binding domain, thus leading to phospho-inositol-glycan (PIG) conjugation after partial cleavage of this terminal sequence. This, in turn, leads to trafficking and positioning of the binding domain on the cell surface. The same cells that took up the first molecule of DNA will express the factor and have increased specific affinity for target DNA that remains extracellular. Cells that did not take up DNA will be at a competitive disadvantage as they will not bear the cell surface target DNA-specific binding domain, which is required for specifically mediated DNA uptake.

Enhanced binding of the target DNA to the target cell will increase the efficiency of DNA internalization and desired intracellular function. This process represents a positive feedback for increased DNA uptake into cells that take up DNA first.

Practical means for determining which transcription factor or combination of factors to use with any particular target DNA

The target DNA, whether a circular or linear plasmid, oligonucleotide, bacterial or mammalian chromosomal fragment, is engineered to bear one or more copies of a DNA recognition sequence for a mammalian or bacterial transcription factor. Many target sequences will already bear one or more such motifs; these can be identified by sequence analysis. Endogenous motifs recognized by these factors also can be identified experimentally by demonstrating that the target DNA binds to one or more of a panel of transcription factors in an appropriate assay format. This provides a practical means for determining which factor or combination of factors to use with any particular target DNA.

Motif(s) in the case of a small oligonucleotide or a DNA plasmid and in the cases where more than one DNA binding protein will be expressed on the cell surface

In the case of a small oligonucleotide or a DNA plasmid (such as used for a DNA vaccine), appropriate motifs can be engineered into the sequence. A particular motif can be engineered in one or more copies, in tandem or dispersed in the target sequence. Alternatively, a set of different motifs can be engineered, in tandem or separated, in cases where more than one DNA binding protein will be expressed on the cell surface.

Evolution Of Bacteriophage Vectors

Using stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly, phage genetics and display technologies to rapidly evolve highly novel, potent, and generic vaccine vehicles

The invention provides methods of obtaining bacteriophage vectors that exhibit desirable properties for use as genetic vaccine vectors. The principle behind the approach provided by the invention is to combine the power of stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly with the extraordinary power of bacteriophage genetics and the wealth of recent advances in phage display technologies to rapidly evolve highly novel, potent, and generic vaccine vehicles.

Methods for delivery of antigens from pathogens to professional APCs, maximizing efficiency through increasing the kinetics and potency of the immune response to the vaccine

The evolved vaccine vehicles can present antigen either (1) in native form on the surface of these APCs for the induction of an antibody response or (2) selectively invade APCs and deliver DNA vaccine constructs to APCs for intracellular expression, processing and presentation to CTLs. More efficient methods for delivery of antigens from pathogens to

professional APCs will increase the kinetics and potency of the immune response to the vaccine.

Affinity maturation process, essential for the generation of antibodies with sufficient affinity to neutralize pathogenic antigens, occurs in germinal centers (spleen) where follicular dendritic cells present protein antigens to B cells and processed antigen fragments to T cells, making efficient delivery of antigens to FDCs essential in increasing the kinetics and potency of the immune response to the immunizing antigen

Genetic vaccine delivery vehicles that are evolved according to the methods of the invention are particularly valuable for the rapid induction of high affinity antibodies which can effectively neutralize viral epitopes or pathogenic toxins such as superantigens or cholera toxin. High affinity antibodies are generated by somatic mutation of low affinity primary response antibodies. This so-called affinity maturation process is essential for the generation of antibodies with sufficient affinity to neutralize pathogenic antigens. Affinity maturation occurs in the spleen in germinal centers where follicular dendritic cells (FDCs), professional antigen presenting cells, present protein antigens to B cells and processed antigen fragments to T cells. Clonally expanding B cell populations which have undergone somatic mutation are selected for those mutant B cells expressing antibodies with improved affinity for antigen. Thus, efficient delivery of antigen to FDCs will increase the kinetics and potency of the immune response to the immunizing antigen. Additionally, processed antigen bound to MHC is required to stimulate antigen specific T cells. Genetic vaccines are particularly efficient at priming class I MHC restricted responses due to intracellular expression of antigen, with a resultant trafficking of antigen fragments to the class I MHC pathway. Thus, invasive bacteriophage vectors capable of delivery of genetic vaccine constructs or protein antigens to FDCs are useful.

Bacteriophage for the purpose of evolution are those that have been genetically well characterized and developed for the display of foreign protein epitopes (of special note was M13 bacteriophage, a small filamentous phage which is a versatile, highly evolvable vehicle for efficient and targeted delivery of protein or DNA vaccine vehicles to cellular targets of interest

Any of several bacteriophage can be evolved according to the methods of the invention. Exemplary bacteriophage for these purposes are those that have been genetically well characterized and developed for the display of foreign protein epitopes; these include,

for example, lambda, T7, and M13 bacteriophage. The filamentous phage M13 is one exemplary vector for use in the methods of the invention. M 13 is a small filamentous bacteriophage that has been used widely to display polypeptide fragments in functional, folded form on the surface of bacteriophage particles. Polypeptides have been fused to both the gene III and gene VIII coat proteins for such display purposes. Thus, M13 is a versatile, highly evolvable vehicle for efficient and targeted delivery of protein or DNA vaccine vehicles to cellular targets of interest.

Improvements in methods (efficient delivery of phage, homing to APCs, and invasion of target cells using experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) bacterial invasion proteins) exemplified for bacteriophage vectors and applicable to other types of genetic vaccine vectors

The following three properties are examples of the type of improvements that can be achieved by use of the methods of the invention to evolve bacteriophage genetic vaccine vectors: (1) efficient delivery of phage to the bloodstream by inhalation or oral delivery, (2) efficient homing to APCs, and (3) efficient invasion of target cells using experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) bacterial invasion proteins. Where M13 is used, fusions can be made to both gene 111 and gene VIII coat proteins so that two evolved properties can be combined into a single phage particle. These studies can be performed in test animals such as laboratory mice so that the evolved constructs can be rapidly characterized with respect to their potency as vaccine vehicles. Evolved inhalable and/or orally deliverable vehicles and evolved invasins will translate directly for use in human cells, while the principles developed in evolving the ability to home to test animal APCs are readily transferable to human cells by performing analogous selections on human APCs. While these methods are exemplified for bacteriophage vectors, the methods are also applicable to other types of genetic vaccine vectors.

Evolution Of Efficient Delivery Of Bacteriophage Vehicles By Inhalation Or Oral Delivery

Method for the formulation of proteins into inhalable colloids that can be absorbed into the blood stream through the lung (preparation involved in the invention)

The invention provides methods for obtaining genetic vaccine vectors that are capable of efficient delivery to the bloodstream upon administration by inhalation or by oral

administration. Methods have been developed for the formulation of proteins into inhalable colloids that can be absorbed into the blood stream through the lung. The mechanisms by which proteins are transported into the blood stream are not clearly understood, and thus improvements are readily approached by evolutionary methods. Using M 13 as an example, the invention involves preparation of a library of, for example, peptide ligands, adhesion molecules, bacterial enterotoxins, and randomly fragmented cDNA, which are fused to gene 111, for example, of M13. Libraries of $>10^{10}$ individual fusions are readily achievable with this technology.

M13 phage enters the blood stream, can be recovered and amplified in E. coli cells, pass through several rounds of enrichment, and be further characterized and evolved by sequencing and reassembling (optionally in combination with other directed evolution methods described herein) the entire phage genome and subjecting the phage to reiterated cycles of delivery, recovery, amplification, and reassembly (optionally in combination with other directed evolution methods described herein)

Screening involves preparation of high titer stocks (e.g., $>10^{12}$ phage particles) in standard colloidal formulations which are delivered intranasally to test animals, such as mice. Blood samples are taken over the course of the ensuing day and circulating phage are amplified in E. coli. It has been established that M13 circulates for long periods in the blood after injection intravenously, and thus it is reasonable to expect that phage that successfully enter the blood stream through the lung can be efficiently recovered and amplified E. coli cells. In one aspect, several rounds of enrichment are applied to the initial libraries in order to enrich for phage that can efficiently enter the blood stream when delivered intranasally. Candidate clones are typically tested individually for their relative efficiency of entry, and the best clones can be further characterized by sequencing to identify the nature of the fusions that confer efficient delivery (of particular interest from the cDNA libraries). Selected clones can be further evolved and for improved entry by reassembling (optionally in combination with other directed evolution methods described herein) the entire phage genome and subjecting the phage to reiterated cycles of delivery, recovery, amplification, and reassembly (optionally in combination with other directed evolution methods described herein).

To obtain vaccine vectors that are effective when taken orally, recombinant vectors prepared through reassembly (optionally in combination with other directed evolution methods described herein) are administered, surviving, stable vectors are recovered from the stomach,

and vectors that efficiently enter the bloodstream and/or lymphatic tissue can be recovered from the blood/lymph.

An analogous procedure is used to obtain vaccine vectors that are effective when delivered orally. A genetic vaccine vector library is prepared by stochastic (e.g. polynucleotide shuffling & interrupted synthesis) and non-stochastic polynucleotide reassembly. The recombinant vectors are packaged and administered to a test animal. Vectors that are stable in the stomach/intestinal environment are recovered, for example, by recovering surviving vectors from the stomach. Vectors that efficiently enter the bloodstream and/or lymphatic tissue can be identified by recovering vectors that reach the blood/lymph. A schematic of this selection method is shown, described &/or referenced herein (including incorporated by reference).

Evolution Of Bacteriophage Vehicles For Efficient Homing To APCs

Two selection formats: the first consisting of enriching the libraries of random peptide ligands and cDNAs used in (A) above for phage which selectively bind to APCs and using either negative or positive selection; the second consists of injecting phage libraries intravenously, collecting target organs of interest, liberating the phage by sonication, further amplifying and enriching.

The invention also provides methods of evolving bacteriophage vectors, as well as other types of genetic vaccine vectors, for efficient homing to professional antigen presenting cells. Libraries of random peptide ligands and cDNAs used in (A) above are enriched for phage which selectively bind to APCs by first negatively selecting for binding to non-APC cell types, and then positively selecting for binding to APCs. The selections is typically performed by mixing high titer stocks of phage from the libraries ($>10^{12}$ phage particles) with cells ($\sim 10^{7}$ cells per selection cycle) and either taking the nonbinding phage (negative selection) or the binding phage from cell pellets (positive selection). An alternative selection format consists of injecting phage libraries intravenously, allowing the libraries to circulate for several hours, collecting target organs of interest (lymph node, spleen), and liberating the phage by sonication. The positively selected phage can be amplified in E. coli and further rounds of enrichment are performed (3- 5 rounds) if further optimization is desired. After the chosen number of rounds, individual phage are characterized for their ability to home to lymphoid organs. The best few candidates can be subjected to further evolution through

iterated rounds of selection, amplification, and reassembly (optionally in combination with other directed evolution methods described herein).

EVOLUTION OF BACTERIOPHAGE FOR INVASION OF APCs

The methods of the invention are also useful for evolving bacteriophage and other genetic vaccine vehicles for invasion of target cells. This opens up the possibility of targeting the class I MHC antigen processing pathways with either internalized protein antigen or antigen expressed by DNA vaccine vehicles carried in by the evolved vector.

Efficient internalization of pathogenic bacteria through invasin interaction with integrins

Invasins comprise a large family of bacterial proteins which interact with integrins and promote the efficient internalization of pathogenic bacteria such as Salmonella.

Reassembly (optionally in combination with other directed evolution methods described herein) of different forms of polynucleotides encoding invasins, cloning as fusions to the M13 gene VIII coat protein gene, preparing libraries and mixing these libraries with target APCs

This embodiment of the invention involves reassembling (optionally in combination with other directed evolution methods described herein) different forms of polynucleotides that encode invasins. For example, two or more genes which encode the invasin family of proteins can be experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis). The experimentally evolved (e.g. by polynucleotide reassembly &/or polynucleotide site-saturation mutagenesis) polynucleotides can be cloned as fusions to the M 13 gene VIII coat protein gene, for example, and high titer stock of such libraries will be prepared. These libraries of bacteriophage can be mixed with target APCs.

Removing free phage and phage bound to the cell surface

After incubation, the cells are exhaustively washed to remove free phage and phage bound to the surface of the cells can be removed by panning against polyclonal anti-M13 antibodies.

Obtaining successful phage, amplifying, reassembling (optionally in combination with other directed evolution methods described herein), and selecting, characterizing for relative invasiveness, combing with gene III fusions (encoding pathogenic epitopes of interest) and testing for relative abilities to induce a CTL response to the pathogenic antigens